PadhAI: 6 Jars of Sigmoid Neuron

One Fourth Labs

Computing Partial Derivatives

How do I compute Δw and Δb

1. Consider the following example
2. Loss = $\frac{1}{5}\Sigma^5_{i=1}(f(x_i) - y_i)^2$ (where f(x_i) refers to the sigmoid function)
3. $\Delta w = \frac{\partial L}{\partial w} = \frac{1}{5}\Sigma^5_{i=1}\frac{\partial}{\partial w}(f(x_i) - y_i)^2$
4. Let's consider only one term in this sum
5. $\nabla w = \frac{\partial}{\partial w}[\frac{1}{2} * (f(x) - y)^2]$ (where ∇w refers to the gradient/partial derivative of L(w,b) w.r.t w)
6. Using chain rule, we expand it
   a. $\nabla w = \frac{1}{2} * [2 * (f(x) - y) * \frac{\partial}{\partial w}(f(x) - y)]$
   b. $\nabla w = (f(x) - y) * \frac{\partial}{\partial w}(f(x))$
   c. $\nabla w = (f(x) - y) * \frac{\partial}{\partial w}(\frac{1}{1 + e^{-(wx+b)}})$, Let's look into the derivative of the sigmoid function in detail
      i. $\frac{\partial}{\partial w}(\frac{1}{1 + e^{-(wx+b)}})$
      ii. $\frac{-1}{(1 + e^{-(wx+b)})^2}\frac{\partial}{\partial w}(e^{-(wx+b)})$
      iii. $\frac{-1}{(1 + e^{-(wx+b)})^2} * (e^{-(wx+b)})\frac{\partial}{\partial w}(-(wx + b))$
      iv. $\frac{-1}{(1 + e^{-(wx+b)})^2} * (e^{-(wx+b)}) * (-x)$
      v. $\frac{1}{(1 + e^{-(wx+b)}} * \frac{(e^{-(wx+b)})}{(1+ e^{-(wx+b)})} * (x)$
      vi. $f(x) * (1 - f(x)) * x$
   d. Therefore, $\nabla w = (f(x) - y) * f(x) * (1 - f(x)) * x$
7. For each of the 5 points
   a. $\Delta w = \frac{1}{5}\Sigma^5_{i=1}(f(x_i) - y_i) * f(x_i) * (1 - f(x_i)) * x_i$
   b. Similarly $\Delta b = \frac{1}{5}\Sigma^5_{i=1}(f(x_i) - y_i) * f(x_i) * (1 - f(x_i))$